DOCUMENT RESUME

ED 213 750 TM 820 164

AUTHOR Lincoln, Yvonna S.; Guba, Egon
TITLE Do Evaluators Wear Grass Skirts? "Going Native" and Ethnocentrism as Problems in Utilization.
PUB DATE Oct 81
NOTE 24p.; Paper presented at the Joint Annual Meeting of the Evaluation Network and the Evaluation Research Society (Austin, TX, September 30 - October 3, 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Anthropology; *Ethnocentrism; *Ethnography; Research Problems; Values
IDENTIFIERS Audits; Debriefing; *Evaluation Problems; *Evaluation Utilization; Journal Writing

ABSTRACT

Anthropological methodology suggests there may be two difficulties with utilizing evaluations over and above the many outlined in several generations of evaluation utilization literature. The first - "going native" - arises when the evaluator has consciously adopted the value and belief system of those in the program he or she is evaluating. The second problem - ethnocentrism - arises when the evaluator has been sufficiently blinded by his or her own value system that he or she fails to see, and thus take into account, divergent values that characterize audiences or recognize how those values reshape objectives or goals. Both types of problems create overt and subtle bias in evaluation reports which in turn leads to skewed judgments, suspect recommendations, loss of evaluator and evaluation credibility, and legitimation of the failure to utilize evaluation results. Solutions to both problems are developed and include, among other strategies, the audit (including means for establishing an audit trail), the reflexive journal, and peer debriefing. (Author/GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED213750

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

DO EVALUATORS WEAR GRASS SKIRTS? "GOING NATIVE" AND ETHNOCENTRISM AS PROBLEMS IN UTILIZATION

by

Yvonna S. Lincoln
University of Kansas

and

Egon Guba
Indiana University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Y. S. Lincoln
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper prepared for presentation at the Joint Annual Meeting of the Evaluation Network and the Evaluation Research Society, Austin, TX, September 30 - October 3, 1981

TM 820 164

Abstract

DO EVALUATORS WEAR GRASS SKIRTS? "GOING NATIVE" AND ETHNOCENTRISM AS PROBLEMS IN UTILIZATION

Anthropological methodology suggests there may be two difficulties with utilizing evaluations over and above the many outlined in several generations of evaluation utilization literature. The first - "going native" - arises when the evaluator has consciously adopted the value and belief system of those in the program he is evaluating. This tendency to be co-opted creates several problems: multiple or competing value systems may not be displayed as clearly or sharply as they should be; results may be skewed toward the favorable, and unfavorable ones may be downplayed or deliberately omitted.

The second problem - ethnocentrism - arises when the evaluator has been sufficiently blinded by his own value system that he fails to see, and thus take into account, divergent values that characterize audiences or recognize how those values reshape objectives or goals. The first instance is a special case of evaluator "corruptibility"; the second, naivete or lack of evaluator introspection.

Both types of problems create overt or subtle bias in evaluation reports which in turn leads to skewed judgments, suspect recommendations, loss of evaluator and evaluation credibility, and legitimation of the failure to utilize evaluation results. The usefulness of such evaluations for policy decisions or analysis is compromised and the likelihood that recommendations that will be utilized is greatly decreased. Neither audiences - whose values are found to be not honored - nor policy shapers - who sense they have incomplete or erroneous information - are moved to act upon reports which embody either of these two faults.

Solutions to both problems are developed and include, among other strategies, the audit (including means for establishing an audit trail), the reflexive journal, and peer debriefing.

As evaluation has come of age as a practice and profession, evaluators and consumers of evaluation studies have become increasingly sophisticated in confronting the problems facing the field. So, while some writers have identified over forty separate evaluation "models" from which a practitioner or consumer might choose, other writers have concerned themselves with barriers to implementing evaluation results and with reasons why the reports of professionals are not more widely utilized. As a result, the "utilization" literature is now nearly as prolific as the theory and model-building literature of yesterday.

Policy analyses, on site interviews and other methods have been employed in the attempt to discover why evaluations are so rarely, or at least incompletely, utilized. Among the reasons cited can be found several categories for failure. The first of those categories might be classified as utilization problems inherent in evaluation design. In this category of utilization failures might be included the issues of relevance, timeliness and validity (Attkisson, Brown and Hargreaves, 1978); poor data management (Niegher, 1979, pp. 125-146); and failure "of the evaluation to conform to criteria for a good evaluation" (Guba, 1975).

The second category might be classified as the evaluation perceptions problem in utilization. In this class of utilization failures resides a series of differing opinions about what evaluation can do, what it is supposed to do, or how it relates to either the program which was designed or

which is currently in operation (and they may not be the same). So, for example, differing "perceptions of the evaluation process and results" held by program evaluators and program managers and funders (Attkisson, Brown and Hargreaves, 1978), "discrepancies between the plans for the program being evaluated and the actual operationalization of those plans" and "the rapidity of change which may render evaluation results obsolete before there is time to use them" (Guba, 1975) may all cause evaluations which conform to perfectly acceptable standards of good practice to be utilized less fully than might have been warranted.

The third cause of failure to utilize might be traced to human and political factors. This set of causes for failure range from the macro level, where "poor organizational response to evaluation findings" can cause difficulty (Attkisson, Brown and Hargreaves, k978) to the personal level. In this last subset would be included inconsistent role demands which interfere with the evaluation process, and unwillingness for one or several persons to "sponsor" or take responsibility for implementing the evaluation results (Neigher, 1979). Finally, is a third subset of the human and political factors (for example, threat level, security of jobs for personnel, and the like), and the corruptibility of evaluators themselves (Guba, 1975, pp. 51-52).

The methodological literature in anthropology suggests that the evaluator may be subject to two difficulties which can result in ignoring, devaluing or underutilizing his

findings. The first, going native, arises when the evalua- has consciously adopted the value and belief system of those in the program being evaluated, principally those of one or more of the relevant audiences. The second difficulty, ethnocentrism, arises when the evaluator has been sufficiently blinded by his own value system that he fails to see, and therefore take in account, the divergent values which characterize his audience(s) or to recognize how those values have reshaped program goals or objectives. The first instance is a special case of evaluator corruptibility (Guba, 1975); the second, an instance of naivete, inexperience, or lack of evaluator introspection. This paper will examine the implications of those two possibilities.

Going Native

When an anthropologis has become so like the group he is studying that he ceases to consider himself a part of the profession - or ceases to consider either his cultural or professional sub-group as his dominant reference group - he is said to have "gone native" (Paul, 1953, p. 438). He quits contributing to the research and begins a "performance - understanding" role (Kolaja, 1956, p. 161) within the studied group. Paul, in a discussion of this problem, named Frank Cushing as an example of an anthropologist who simply refused to continue publishing the results of his field studies. Identification with the "natives", or co-optation, as a persistent problem of inquirer identification, has been a part of the "warnings and advice" given to new participant observers

for several decades. Gold (1969) suggests that going native is almost always a result of naivete, and happens as an unfortunate accident. In the process of attempting to gain Verstehen, he asserts, "... the field worker may over-identify with the informant and start to lose his research perspective by 'going native'" (p.36). Moreover, "prolonged direct participation entails the risk that the researcher will lose his detached wonder and fail to discover certain phenomena that the relatively uninvolved researcher would discover" (p. 63-54, italics added). Gold has most assuredly given anthropologists the benefit of the doubt, but it is not certain that modern evaluators (save perhaps the youngest and most inexperienced) would happen into "going native" because of over-identification, and one wonders whether the benefit of the doubt is appropriate for them.

The problem of going native has some freshness for evaluators. There are many social action programs which address the pressing needs of virtually disenfranchised groups of citizens, and every egalitarian instinct causes one to wish them success and to want to do something which will ensure that success. Humanitarian urges cause us to minimize negative consequences in favor of clearly positive outcomes. Multiple or competing value systems may not be displayed and contrasted as clearly or sharply as they should be to have the evaluation considered fair; results may be skewed toward

the favorable or positive; and unfavorable effects of side-effects may be downplayed or deliberately omitted.

But allowing oneself to be co-opted is likely to result in distorting or overlooking two vital areas: (1) the true needs of the "natives", or program audience(s), for worthwhile programs, and (2) the broader social needs for discriminating but fair decisions to be made from among a variety of potentially worthwhile but resource-scarce programs. Thus, while the evaluator may feel he is serving a larger social purpose by attending to positive aspects of the program and minimizing negative aspects, he may actually be perpetrating a disservice to both program targets and society at large. The boundaries of his faithlessness to the task at hand--sound evaluation with discriminating choices displayed for all groups to see--are sufficiently wide to cause disrepute to settle on the evaluation community as a whole. Indeed, commented Radin, "True participation is simply out of the question and romantic participation obscures the situation completely. For any ethnologist to imagine that anything can be gained by 'going native' is a delusion and a snare" (1933, p. 227, italics added).

Ethnocentrism

Gold, in discussing problems in collecting and analyzing field data, comments that there is a flip side to the "going native" coin:

> "... with respect to achieving rapport in a field relationship, ethnocentrism may be considered the logical opposite of "going native". Ethnocentrism occurs whenever a field worker cannot or will not interact meaningfully with an informant ... " (1969, p. 37)

Vidich confirms enthnocentrism as a legitimate influence in the failure to collect and interpret meaningful data, and comments that is it particularly a problem when a field worker (read: evaluator) enters into a setting in his own culture:

> "This naive attitude cannot be assumed /of being a stranger to one's own culture/ in working in his own culture for the simple reason that the respondent cannot accept it as plausible. In fact, the difficulty of securing data may be increased by the 'ethnocentrism' of some respondents who assume that their own experiences are similar to those of others." (p. 82)

Gold and Vidich imply that ethnocentrism works in both directions: both on the part of the evaluator and on the part of the respondents, who may feel that the evaluator shares their experiences and therefore has a common and shared set of values and beliefs. While neither instance services evaluation needs, ethnocentrism on the part of the evaluator is the more crucial, since it may cause him to miss components of cultures or attitudes which are crucial to functioning of programs.

The problem of ethnocentrism causes particular difficulty when the evaluator must deal with pluralistic value contexts. Then one of his principal roles is to discover, explicate and contrast the competing value sets held by various audiences. Another important task is to determine the consequences of pursuing any and all given value structures in order to reflect their implications for program direction and operation. To the extent that he is incapacitated by this ethnocentrism in determining structures exist in a given program setting, he

has served the evaluation effort less fully than he might.

How Serious Are The Problems?

These warnings from the anthropological literature are really somewhat more serious than they appear to be. As part of the earliest methodological literature in the field, they deal with one of the more serious potential flaws to field work (and by extension, to evaluation), bias. Bias is one of the more frequent charges leveled at evaluations, and one of the more subtle of the political factors which militate against using evaluation results. Bias is an especially potent charge when the subject is naturalistic evaluations, which rely heavily on the identification, comparison and interpretation of competing value structures in program sites. The kinds of research and evaluation which the naturalist carries out lean so strongly on the human as instrument and his expertise and judgment in discovering values and displaying them in contrasting form that both audiences for the evaluation within the program and critics of the evaluation (or program) without may deliberately choose to ignore reports which they can see conveniently labeled as biased.

Some Responses

How does one go about avoiding these pitfalls to evaluation? Both going native and ethnocentrism create overt or subtle bias in evaluation reports which in turn leads to skewed judgments, suspect recommendations, loss of evaluator and evaluation credibility, and legitimation of the failure

to utilize evaluation results. The usefulness of such evaluations for policy decisions or analysis is compromised and the likelihood that recommendations will be utilized is greatly decreased. Neither audiences--whose values are found to be dishonored, or not honored at all--nor policy makers and shapers--who sense they have incomplete or erroneous information--are moved to act upon reports which embody either of these two faults.

The question of possible responses is twofold: on the one hand, how can an evaluator avoid one or both of these faults and prevent their occurrence in himself; on the other hand, how can those who receive such reports be certain they have a report which is free from the same biases? The answer to the questions involves three strategies--the reflexive journal, peer debriefing, and auditing--which help the evaluator to develop self-awareness and help others to "track" the development of his insights, data and interpretations. These strategies increase the possibility of utilization of evaluation results since they mean more publicly inspectable results and invite comparison with other, equally public interpretations. Each will be discussed in turn.

Strategies for Improving the Changes of Utilization of Evaluation

The Reflexive Journal. Progoff (1975) has pointed out that journals have purposes much larger than merely keeping track of one's day-to-day activities. They can, in fact, provide major thematic analysis units for exploring one's "life scripts" and for re-writing and re-thinking varied aspects of those

scripts which one may find dissatisfying, pain-producing or unfulfilling. In addition, such a journal can offer retrospective evidence about changes in attitudes, values and beliefs in the writer (i.e., the evaluator) over time. Likewise, Reinharz (1978) has indicated that reports of investigations typically describe the problem and the method; and presumably, if one knows problem and method, one can then tell whether outcomes and conclusions are trustworthy, rigorous, and/or relevant. But, she counters, research represents an integration not only of problem and method, but also of the person who is doing it. To some extent, Reinharz's book (1978) is a retrospective reflexive journal, as are certain writings of Wax (1971), Mead (1972), and Zigarmi and Zigarmi (1978). The effort to understand how the inquirer changed as a result of his own inquiries, or, as Wax described it, "... the things I learned" (1971, p. 363), is a perennial problem for the reflective observer, and more so because of the nature of some forms of inquiry.

In the case of naturalistic evaluation in particular, designs tend to be emergent; that is, what happens on any given day depends on what has happened on all the days which came before. It is essential that there be some means for continuously evolving, summarizing, and projecting, to provide for orderly emergence and evolution of the design and data collection efforts.

It is also the case in naturalistic evaluations (those which do not rely on preordinate and fixed designs which prespecify each step in the evaluation process) that one good method for establishing at least some aspects of trustworthiness (or rigor) is the audit. The reflexive fieldwork journals of the evaluator provide a rich resource in the audit trail.

To be reflexive means to turn back in on; thus the reflesive journal turns back on the investigatbr or evaluator and reflects him to himself (and, after the evaluation or inquiry, to others). Not only is it then clear how the design of the evaluation is emerging, but it also becomes clear how the evaluator's qualities as a "smart" instrument (and inquirer) are also emerging.

A "relatively complete" reflexive journal probably contains the following five sections:

1. A log of evolving perceptions. This log begins with the written perceptions of the evaluator prior to his entrance on site; and additional statements are written at regular intervals. The separate renderings may be kept in escrow by someone else--usually an auditor or peer debriefer--who is not directly connected with the program or project evaluation. At the end of the project (or during its duration if it is of sufficient time lapse), these perceptual statements may be compared to one another and/or to aspects of the design and/or analysis to check whether learning is occuring, whether original perceptions and beliefs perseverate, whether later findings

are clouded by these perceptions, and whether or not the evaluator has "learned" anything "new" or changed himself.

2. A log of day-to-day procedures. The purpose of this log is to indicate in diary or chronographic form exactly what was accomplished every day. The most important use of this component of the journal will be in the evaluation audit process.

3. A log of methodological decision points. This particular log of the reflexive journal is entered on an ad hoc basis as necessary, and should record all major methodological decisions, such as explicating new design steps, decisions on instrumentation, finalization of an analysis category set or the like, and such decisions should be entered together with reasons or rationale for the action taken or decision made. The auditor is the chief audience for this log.

4. A log of day-to-day personal introspections. Here one lays out in diary form one's own thoughts and feelings, including stresses one is almost bound to undergo (Zigarmi and Zigarmi, 1978) and frustrations one encounters, and how those feelings and situations are perceived to be changing. It is important to have some insight into one's self and to work on generaring that insight. The chief use to which this log can be put is to test for bias in the evaluator and to relate decisions about design and procedures to it later.

5. A log of developing insights and hypotheses. The emergent design will depend almost exclusively on how the evaluator takes advantage of what he has already learned. The purpose of this section is to keep readily available an up-to-date summary of where one is with respect to knowledge of the situation and working hypotheses about it. Those working hypotheses which have been discarded or "outgrown" ought to be relegated to historical files on the project, and retained as part of the audit trail. There are several uses for this log, including the guiding of the inquiry, shaping of the emergent design, providing the basis for subsequent data collection and analyzing activities, and for post hoc auditing procedures.

One of the reasons so few such "journals" are kept is that the keeping of one requires enormous discipline for an evaluator, especially when he knows that his entries will be open to inspection, for instance, by debriefers and perhaps later by auditors. The logs provide an ongoing record of the plans, operations, decisions, travails and other mental processes and activities of the evaluator.

Peer De-briefing. This particular strategy for maintaining one's sense of reality counteracts both the process of "going native" and ethnocentrism, but it also contributes a third benefit to the ongoing evaluation: it allows for working hypotheses to be tested by a peer or colleague who may or may not be involved with the evaluation firsthand. Peer

de-briefing can take place either with a person or persons outside and removed from the evaluation effort, or with members of a team of evaluators who are operating in a split-half or teaming mode, collecting independent sets of data, then reviewing those data for discovery and verification of meanings and findings. In either instance, the purpose of such de-briefings is to keep the evaluator--and his data collection and interpretation procedures--"on track," usually by reflecting on a priori written statements of the evaluator, by reviewing the reflexive journal and by careful questioning of the data which are reported to the debriefer.

There are three reasons why an evaluator might want to engage a peer de-briefer. The first is relief from the pressures of long and hard days of trying to "fit" or coping with long and complicated interviews which must then be transcribed, fleshed out and completed late at night and alone, or the strains of observation, especially when one is not certain what one is observing, make for the necessity of having someone simply to talk to (better yet, to act as a formal debriefer) (Zigarmi and Zigarmi, 1978).

The second is to recover from what Wax calls "immersion," by which is meant"...a joint process (of becoming a member of a society or culture of living people)...involving numerous accommodations and adjustments by both the fieldworker and the people who 'accept' him" (1971, p. 43, italics the author's own). She describes forms of immersion (which occasionally

lead to "going native" or otherwise losing perspective on the research):

> "For example, a fieldworker may become so fascinated by the new, exciting, and significant things he is learning, that he may spend months passionately and persistently thinking of nothing else. Simultaneously, he will find himself becoming personally or socially involved in the community, not only because of his developing relationships with acquaintances, employees, and friends, but because, to some degree at least, he is now really beginning to lose touch with his own people and with the world outside...
>
> There is another and deeper kind of immersion which may occur after a fieldworker has truly become involved with the 'living people' in the society he is studying. Indeed, he may be unaware that he is 'immersed' until he is given the opportunity to leave the field for a pleasant vacation and finds that he does not want to go. Sometimes, his new and hard-won social ties and relations may mean, or seem to mean, more to him than his ties with his own people." (1971, pp.43-44)

Thus, even reasonable people may disagree on what the problem is, but the need for getting away from the site, and for having someone (a peer) debrief the inquirer on his research is imperative.

A third reason to utilize peer debriefing is to provide for an external check. Douglas reports that, when operating in a team situation (more than one researcher or evaluator), "...team members do not merely provide support. They also provide a vital check on each other. Field researchers have always bemoaned the almost total lack of internal checks and of retesting their studies. The few instances of retesting that have been done ... have shown that they need it badly.

Retests tend to show that each participant observer has gotten at different parts of given different interpretations of things". (1976, p. 217, italics ours).

The Audit and Audit Trail. Perhaps the best method for insuring that evaluators (and indeed, all field-oriented naturalistic inquirers) keep themselves--and are kept--honest, is the establishment of an audit trail. Not only does such a "track" prevent the evaluator from becoming so enmeshed in his own values that he fails to discover and portray multiple and competing value systems in the context, it also allows others to discover when he has become so enmeshed in the site and the program that he has "gone native" or remained ethnocentric and betrayed his real purposes.

The audit is a major technique modelled on the fiscal audit which can be used both for dependability (analogous to reliability) and confirmability (analogous to objectivity) purposes (Ridings, 1980; Guba, 1981; Halpern, 1981). That the techniques will suffice for these purposes is well illustrated be a comment made by Cronbach and Suppes in describing a feature that distinguishes disciplined inquiry from other forms:

> "...the report of a disciplined inquiry has a texture that displays the raw materials entering the argument and the logical processes by which they were compressed and rearranged to make the conclusion credible." (1969, p. 16)

The audit is a means for carrying out this kind of public examination. It is based on the metaphor of the fiscal auditor, who, when called in to examine the books of a client corporation or business, has two major tasks that he is expected to accomplish:

1. To examine the processes by which the local accounts are kept, not so much to assure that there has been no fraud (although the detection of fraud is one of the auditor's functions) than to assure that the books represent a fair statement of the company's fiscal position. The auditor is particularly concerned that there has been no "creative accounting", as it is called in the trade, for example, to make the company look more attractive to stock purchasers or to possible buyers. The auditor's major taske is that he be able to certify that the processes used by the local accountants fall within the bounds of acceptable professional practice.

2. To examine the products of the local accounting processes, to ascertain that every entry in the books can be authenticated either by direct documentation or by solicited confirming statements, that the "bottom line" is correct, and that interpretations made of the accounts in any fiscal statement based on them are appropriate.

In similar fashion, the inquiry or evaluation auditor has two tasks: to review the inquiry processes to be certain that they fall within the bounds of acceptable professional practice (norms), and to review the inquiry products to be certain that they can be substantiated from the data collected (a step, by the way, which is almost never applied systematically in conventional research). The former task is equivalent to establishing the dependability of the inquiry and the latter its confirmability

(of both data and conclusions). The issue is not whether the evaluator has carried out the processes or reached the conclusions in the same way the auditor would have done, but whether the evaluator has carried out the processes in a reasonable manner. Thus, incidentally, replication is not called for as a criterion, but rationality is.

The actions of the auditor differ depending on whether he is doing a dependability audit or a confirmability audit, although of course he may do both. In the role of a dependability auditor, he will: examine all of the documentation from the point of view of its acceptability within the norms of good naturalistic professional practice, and; certify that the inquiry has been adequately and fairly carried out and attest to that fact with a formal statement. In the role of a confirmability auditor, he will: examine all analyzed data; and compare a sample of analyzed data with original data items such as interview notes or documents to satisfy himself that: the data items have been reasonable unitized, and categorized, that individual data items have been reasonable assigned to appropriate taxons or categories; that conclusions are documented in terms of the category system; and that conclusions are triangulated with respect to multiple data sources. In addition, he will certify that the inquiry products are properly founded on the data and reasonably interpreted from them, and attest to that fact with a formal statement.

If auditing is to be done properly, the auditor should be involved form the beginning to be certain that a proper audit trail is being left. The auditor may wish to work with the investigator

throughout the study, although of course the major work of the auditor comes post facto.

Summary

We have tried to indicate that there are several types of problems which creep into evaluation reporting which affect the report's potential for utilization. Those problems are suggested by the anthropology methodological literature in anthropology dating back nearly 50 years, and revolve about either the evaluator's "going native"-- or consciously or unconsiously adopting the value structure of his respondents to the extent that he ceases to relate to his professional peer group as a professional -- and the inverse of this situation, ethnocentrism -- failing to recognize, portray and honor the multiple value perspectives he might find in a context because his own value orientations may have blinded him to other perspectives.

Either of these two forms of bias is likely to result in failure to utilize the results of evaluation; in alienation of program targets; who realize their value positions are not honored; in the inability of such reports to be employed in policy analyses; and in a loss of evaluator and evaluation credibility.

We have suggested three strategies which help detect and counteract bias in evaluation reporting. The first, the reflexive journal, allows the evaluator to chart his own growth and to mark where he is not "learning" as program audiences are teaching him, or where he may be unconscious prejudices which are not originally known to him. The second, peer debriefing, allows a colleague to provide this check, both on the evaluator's progress as a "smart" instrument, and to verify that the study is proceeding in an appro-

priate collection and analytic mode. The third, the evaluation audit, provides for a methodological and analytic accounting by an outside party to the evaluation, in order to insure that proper procedures, both with respect to data collection and with respect to data analysis, have been carried out. In addition, the audit provides a means for suggesting proper "audit trails", or accounts of the research and decision points, which will extend the means for peer reviewers to certify that the research has proceeded along standard and accepted canons of good practice.

Bibliography

Attkisson, C.C., W.A. Hargreaves, M.J. Horowitz, and J.E. Sorenson, Eds. Evaluation of Human Service Programs, New York: Academic Press, 1978.

Cronbach, Lee J and Suppes, Patrick, Eds. Research for Tomorrow's Schools: Disciplined Inquiry for Education. Report of the National Academy of Education. New York: MacMillan and Co., 1969.

Douglas, Jack D., Investigative Social Research. Beverly Hills, California: Page Publications, 1976.

Gold, Raymond L. "Roles in Sociological Field Observations", in McCall and Simmons, Eds. Issues in Participant Observation: A Text and Reader. Reading, Massachusetts: Addison-Wesley Publishing, 1969.

Guba, Egon G. "Criteria for Assessing the Trustworthiness of Naturalistic Inquiries", Educational Communications and Technology Journal, 29 (Summer, 1981), pp.75-91.

Guba, Egon G. "Problems in Utilizing the Results of Evaluation", Journal of Research and Development in Education, 8 (1975, Spring) pp. 42-54.

Halpern, Edward. Evaluating Non-quantitative Inquiries: A Methodological Exploration. Unpublished doctoral qualifying paper, Indiana University, 1981.

Kolaja, Jiri, "A Contribution to the Theory of Participant Observation," Social Forces, 35 (1956) pp. 159-163.

Mead, Margaret. Blackberry Winter: My Earlier Years. New York: Morrow, 1972.

Neigher, William D. "Effort and Influence: Enhancing the Impact of Mental Health Program Evaluation at the Agency Level", in Robert F. Rich, Ed. Translating Evaluation Into Policy. Beverly Hills, California: Sage Publications, Inc., 1979.

Patton, Michael Q. Utilization - Focused Evaluation. Beverly Hills, California: Sage Publications, 1979.

Paul, Benjamin D., "Interview Technique and Field Relationships," in A. L. Kroeber (Ed.) Anthropology Today. Chicago: University of Chicago Press, 1953.

Progoff, Ira. At a Journal Workshop. New York: Dialogue House Library, 1975.

Radin, Paul. The Method and Theory of Ethnology. New York: McGraw-Hill Book Co., Inc., 1933.

Reinharz, Shulamit. On Becoming A Social Scientist. San Francisco: Jossey - Bass Publisher, 1978.

Ridings, Jeri M. Standard Setting in Accounting and Auditing: Considerations for Educational Evaluation. Unpublished doctoral dissertation, Western Michigan University, Kalamazoo, Michigan, 1980.

Smith, Nick L., Ed. Metaphors for Evaluation: Sources of New Methods. Beverly Hills, California: Sage Publishing Co., 1981.

Smith, Nick L., Ed. New Techniques for Evaluation. Beverly Hills, California: Sage Publishing Co., 1981.

Spradley, James P. The Ethnographic Interview. New York: Holt, Rinehart and Winston, 1979.

Tobier, A., Evaluation Reconsidered. Paper read to the American Educational Research Association, Chicago, 1972. ERIC No. ED 080 453.

Vidich, Arthur J., "Participant Observation and the Collection and Interpretation of Data." In McCall, G.J. and Simmons, J.L., Eds. Issues in Participant Observation: A Text and Reader. Reading, Massachusetts: Addison - Wesley Publishing Co., 1969.

Wax, Rosalie H., Doing Fieldwork: Warnings and Advice. Chicago: Univeristy of Chicago Press, 1971.

Zigarmi, D. and Zigarmi, P. The Psychological Stresses of Ethnographic Research. Paper read at the American Educational Research Association, Toronto, 1978.